



# Virtual Drug Screening (VDS) Stream:

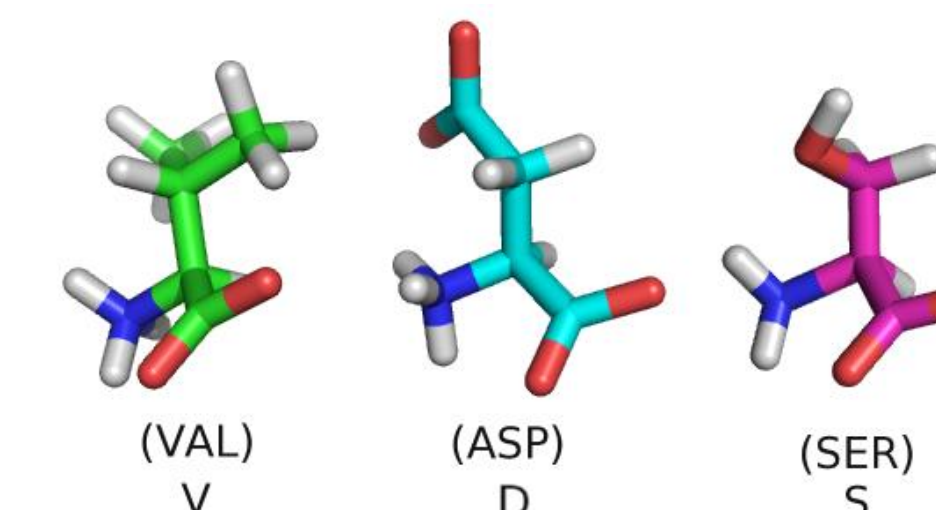
## Can new drugs be identified from virtual libraries of molecules?

**Students:** Ariel Chen, Alice Chen, Antonio Garcia, Aaron Hong, Ashlee Wolf, Alyssa Kahl, Alakh Rajan, Aditi Singh, Anita Vasudevan, Brendan Chou, Brandy Couchot, Caroline Chen, D'ondria Peters, Daniel Dzurko, Emily Johnson, Feng Gao, Grayson Northcutt, Grant Tippet, Grace Truong, Gautam Warriar, Hyun-Young Lee, Imran Zafar, Jacqueline Elizondo, Jensen Gary, Julia Chernis, Jesus De La O, James Tran, Jessica No, Jiaqi Zhang, Karuna Antani, Khady Diack, Renee French, Kevin Elwood, Katherine Vuong, Keren Kahsai, Kavya Kolavasi, Karinna Luis, Kelsey Cantu, Keely White, Miguel Torre, Marianna Uriegas, Manuel Zamora, Madeline Jiang, Mohammad Qasim Adil, Min Seon, Muhammad Akhtar, Melissa Huynh, Nicolet Finger, Nancy Ibrahim, Nicole Welch, Priya Patel, Ramiro Rodriguez, Shivani Bindal, Stella Miller, Dax Fisher-Garibay, Shannon Richard, Selena Ilbeig, Seo Kim, Shawn Okpara, So-Youn Park, Spencer Coronado, Serena Zadoo, Hunter Torres, Tracy Cheng, Victoria Gomez, William Ip, William Eveland

**Mentors:** Suman Augsteen, Andrew Butler, Aldo Del Olmo, Divya Goruganthu, Janice Kim, Max Magallanes, Urvashi Mitbander, Brandon Nguyen, Daniel O'Brien, Kaarthik Rajendran, Michael Tra, Ruifei Wang, Sadhana Balasubramanyam, Yunle Huang, Jose Olmos, Ruoyi Pu

**Teaching Assistant:** Thao Pham **Instructors:** Josh Beckham (RE), Jon Robertus (PI)

Freshman Research Initiative, Dept. of Chemistry and Biochemistry, College of Natural Sciences, The University of Texas at Austin



### ABSTRACT

Enzymatic proteins are at the heart of many disease processes. The ability to effectively target and inhibit their molecular function provides an opportunity to mitigate the deleterious outcomes of disease states in humans. However, **identifying new drug leads using traditional methods is an expensive and time consuming process.** This research stream uses **computers to sift through libraries of chemical structures** and predict which ones may bind most effectively to a protein that is a potential drug target. Through this work students are introduced to fundamental features of protein structures and of protein-ligand interactions. Virtual drug screening software is used and the results are then visualized and interpreted with a molecular graphics program. **Several of the best candidate molecules can then be tested in the wet lab to determine their efficacy in comparison to the computational predictions.**

### INTRODUCTION

#### EXAMPLE DISEASE

**Cancer is the second most common cause of death in the US behind heart disease<sup>1</sup>.** Each type of cancer is essentially a unique disease due to the variety of pathways that are activated or suppressed to bring about aberrant cell growth and functionality. Consequently targeting each individual cancer separately is a daunting task. However, one approach is to inhibit a molecular pathway that is particularly common in all cancers.

#### MOLECULAR TARGET

**Human dihydrofolate reductase (DHFR)** was chosen as the drug target for our virtual screening. **This protein acts as an enzyme which facilitates the production of tetrahydrofolic acid from the substrate dihydrofolate** (see Fig 1). Tetrahydrofolate is required by rapidly dividing cells to synthesize DNA precursors. Since cancer cells divide much more rapidly than ordinary cells, inhibition of DHFR is a useful therapy for many types of cancer. In this research, we sought to compare the binding of 15 different ligands to the DHFR enzyme.

### METHODS

#### MOLECULAR VISUALIZATION

An open source software program called PyMol<sup>2</sup> was used in order to view the structure of the DHFR enzyme and the candidate ligands. The three dimensional structure files (.pdb), determined by X-ray crystallography, were obtained from the literature at the Protein Data Bank ([www.rcsb.org/pdb](http://www.rcsb.org/pdb))<sup>3</sup>. The images were manipulated to study the primary, secondary and tertiary structure of the protein (Fig. 4). Then the binding site was analyzed to determine the most likely amino acids to be involved in protein-ligand interactions.

### LIGAND LIBRARY

A small electronic library of the 3-dimensional structures for a set of ligands numbered 1-15 was used for docking into the active site of DHFR (Fig. 2 with corresponding names shown in Table 1). These ligands were chosen from the literature for their known interactions with the enzyme and their similarity to the natural substrate, dihydrofolate (shown in Fig. 1). The structural similarities can be seen by the presence of two- membered rings similar to the two membered pterin ring system in dihydrofolate.

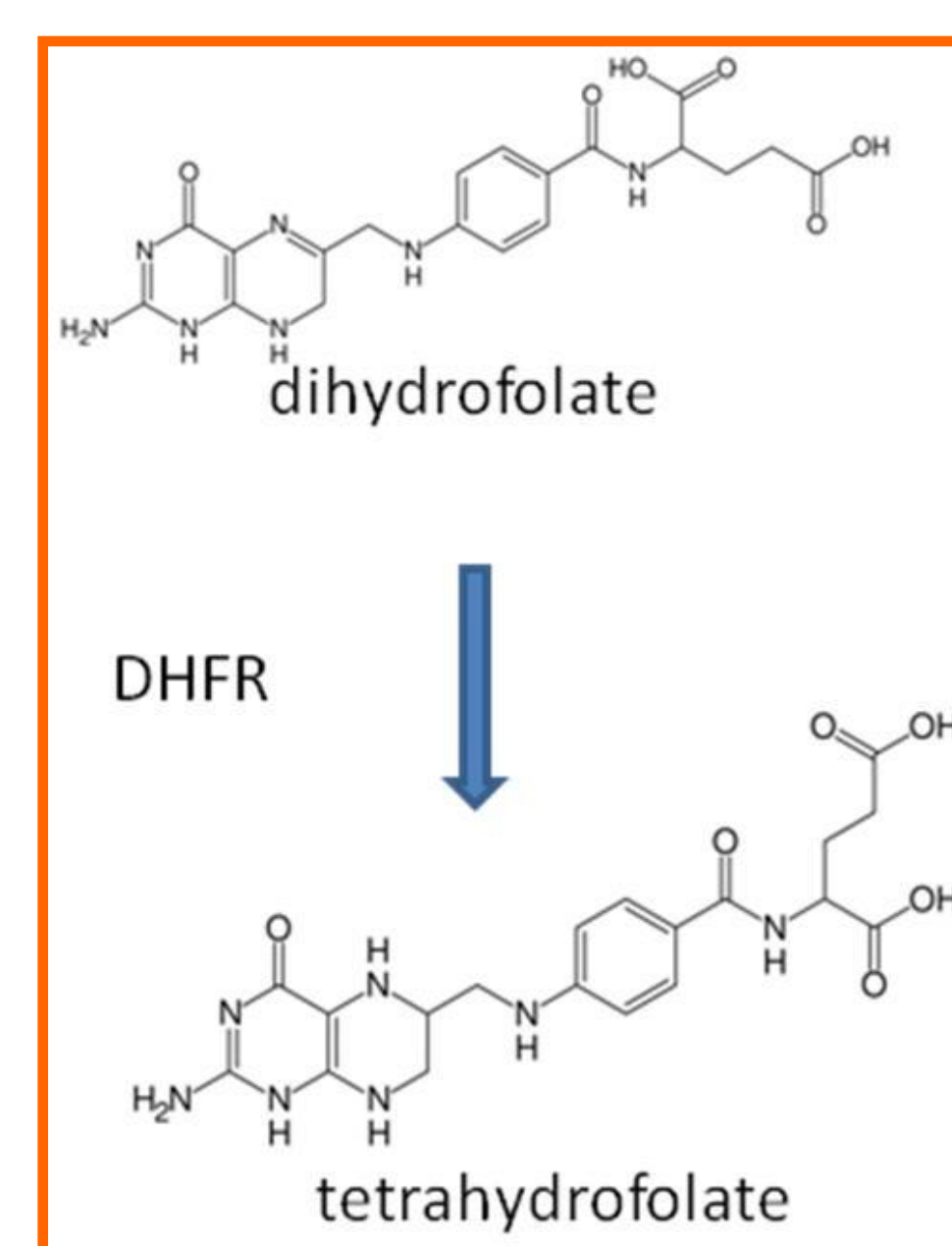
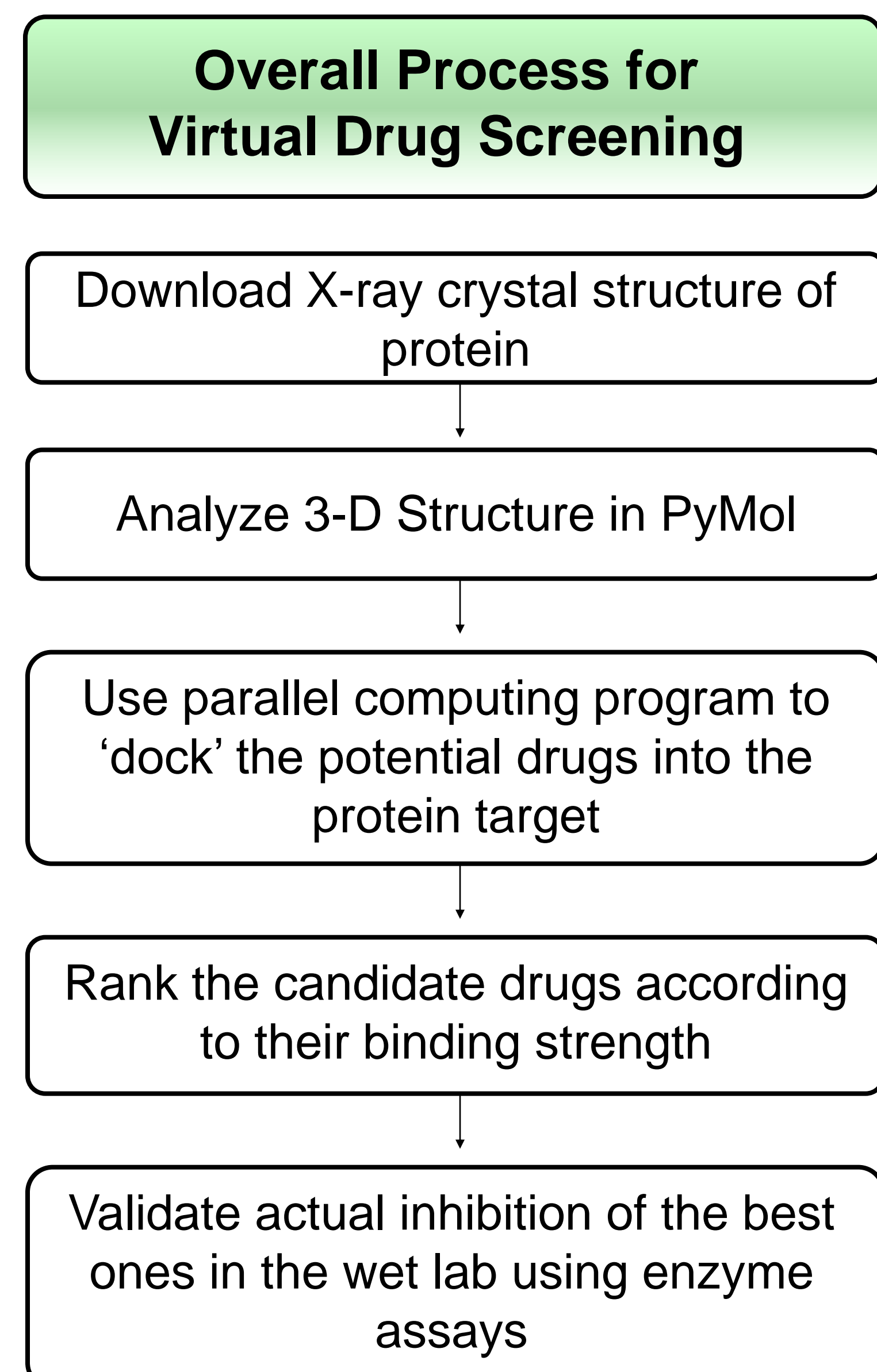


Fig 1. Dihydrofolate Reductase catalyzes the reduction of dihydrofolate to tetrahydrofolate – a necessary step for the synthesis of DNA

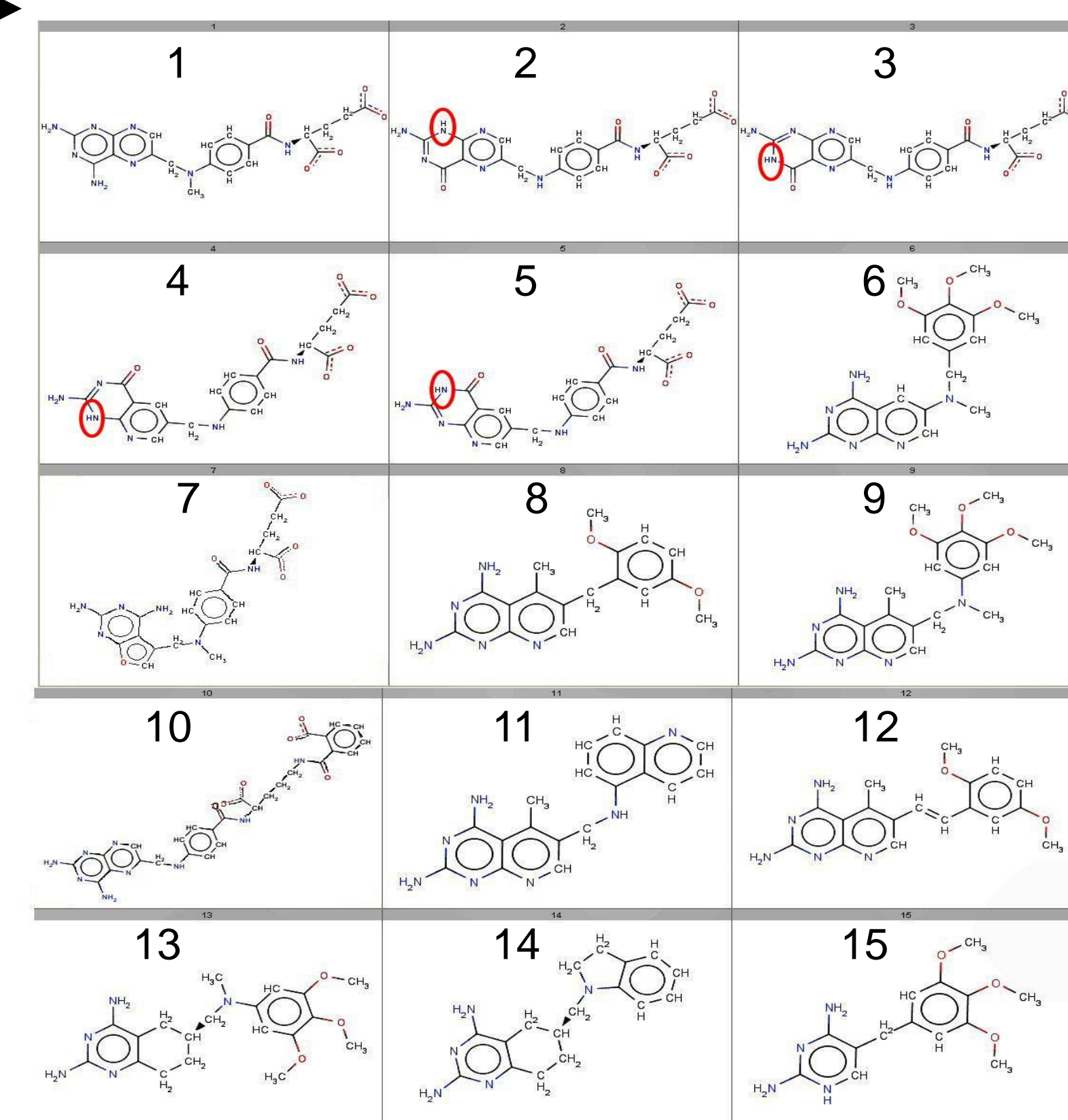
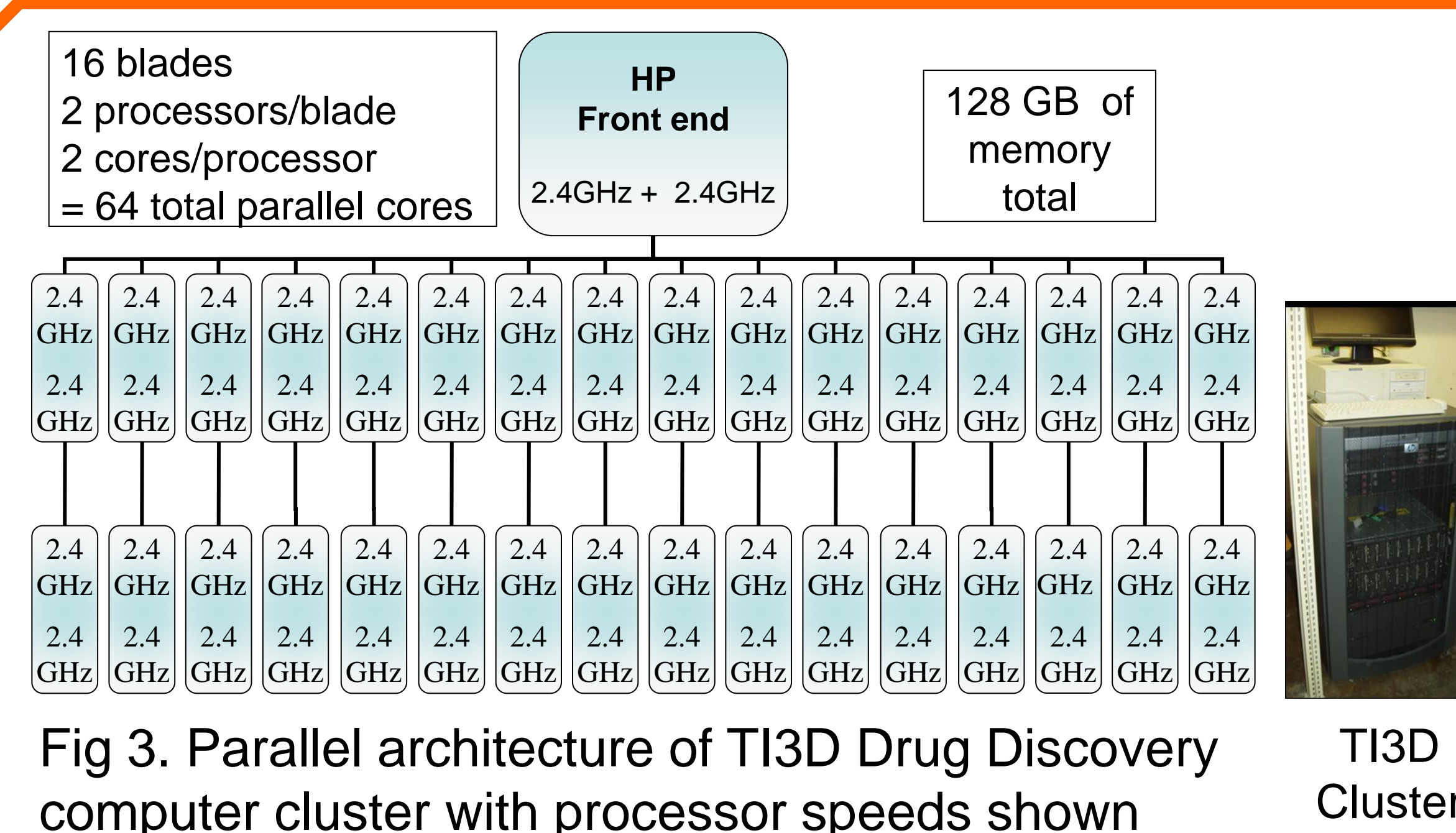


Fig 2. Ligands (#1-15) used for virtual drug screening against DHFR enzyme



### METHODS

#### VIRTUAL SCREENING

In order to dock each ligand into the three dimensional model of the protein target, a software program called GOLD<sup>4</sup> is implemented in a parallel architecture on the TI3D Drug Discovery Cluster, which contains 16 HP Proliant BL35P blade servers, each with 2 dual core AMD Opteron 2.4 GHz processors for a total of 64 processors (see Fig 3). Each blade contains 8 GB of memory and a 6 GB ATA hard disk drive. The front-end of the cluster is an HP xw9300 Workstation, equipped with a dual core AMD Opteron 2.4 GHz processor, an NVIDIA Quadro FX4500 graphics card, 4 GB memory, and two 500 GB SATA hard drives that is backed up through a RAID5 network.

The GOLD application is capable of scanning through the many different conformations of each ligand that are due to rotatable bonds. For each corresponding docking conformation a Fitness Score is assigned which quantifies the relative strength of the binding of the ligand to the active site of the protein target. This score is an aggregate which incorporates the change in free energies of the binding due to several different chemical interactions: Van der Waals forces, hydrogen bonds, hydrophobic interactions and strain penalties (see Eqn. 1 below).

$$\Delta G_{bind} = \Delta G_{vdw} + \Delta G_{H-bond} + \Delta G_{hydrophobic} + \Delta G_{rotor} + \Delta G_0 \quad \text{Eqn. 1}$$

### RESULTS

The initial computational screening using GOLD software found the best binding conformation out of 10 different ones for each ligand listed below and gave it a Fitness Score (Table 1). **A higher score corresponds to stronger binding between the enzyme and ligand and, therefore, greater predicted drug activity.** Notice that the natural substrates (the folic acids – FOL and FOLH3) have the highest scores. The anti-cancer drug methotrexate (MTX) scores relatively high, while the anti-bacterial drug trimethoprim (TRR) scores low. Figure 4 shows the representation of the binding of MTX and DHFR as shown in PyMol with orange lines depicting bonds.

Ligand #	Fitness Score	Ligand Identifier	Full Ligand Name
2	111.05	FOL	FOLIC ACID - N1H tautomer
3	102.45	FOLH3	FOLIC ACID - N3H tautomer
5	101.19	DZFN3	5-DEAZOFOLATE - N3H tautomer
10	96.33	COP	NA-(4-AMINO-4-DEOXYPTEROYL)-ND-HEMIPHTHALOYL-L-ORNITHINE
4	91.43	DZF	5-DEAZOFOLATE - N1 tautomer
7	88.91	MOT	FUOPYRIMIDINE
1	87.12	MTX	METHOTREXATE
14	70.85	TQT	6-(OCTAHYDRO-1H-INDOL-1-YL METHYL)DECAHYDROQUINAZOLINE-2,4-DIAMINE
6	66.44	DTM	Pyrido[2,3-d]pyrimidine N9-C10
13	66.32	TQD	6-[(METHYL(3,4,5-TRIMETHOXYCYCLOHEXYL)AMINO)METHYL]DECAHYDROQUINAZOLINE-2,4-DIAMINE
15	65.7	TRR	TRIMETHOPRIM
11	37.1	LIH	6-[(5-QUINOLYLAMINO)METHYL]-2,4-DIAMINO-5-METHYLPYRIDO[2,3-D]PYRIMIDINE
9	36.48	CO4	5-Methyl-6-N-Methylanilino Pyridopyrimidine
12	35.21	LII	6-(2-[2,5-DIMETHOXYPHENYL]ETHEN-1-YL)-2,4-DIAMINO-5-METHYLPYRIDO[2,3-D]PYRIMIDINE
8	34.04	MXA	6-(2,5-DIMETHOXY-BENZYL)-5-METHYL-PYRIDO[2,3-D]PYRIMIDINE-2,4-DIAMINE

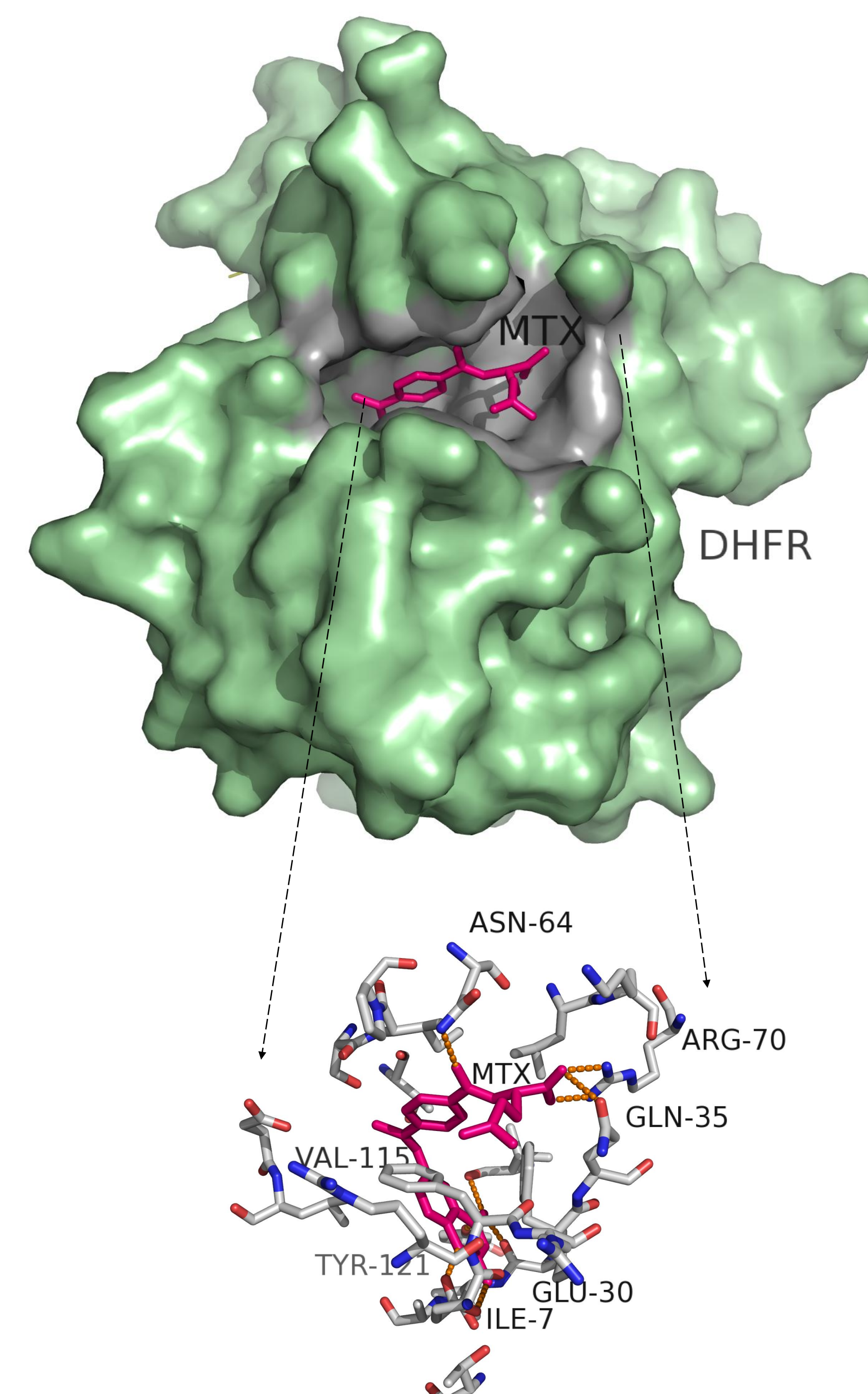


Fig 4. Binding of the ligand, MTX, (colored magenta) in the active site of the enzyme dihydrofolate reductase (DHFR) with amino acids labeled that are involved in binding. Carbon (gray), oxygen (red), nitrogen (blue). Orange lines are bonds.



Researchers of the Virtual Drug Screening class



This work is funded by the National Science Foundation and the Howard Hughes Medical Institute

### CONCLUSIONS

While not providing proof that any compound will inhibit a given enzyme *in vivo*, virtual drug screening does help the researcher to **narrow down the field of best possible ligands that could then be tested in the lab**, thereby reducing the time and cost of bringing new drugs to market.

### FUTURE DIRECTIONS (Summer & Fall)

Using the skills from this research stream, students will select other protein targets for virtual screening, such as enzymes involved in bacterial and fungal infections or those used in bioterrorism (like anthrax or ricin). After screening for the best predicted binding, the actual compounds will be obtained commercially for validation in the wet lab using light absorbance based enzymatic assays that quantify the level of inhibition for each drug candidate.

#### REFERENCES

- Center for Disease Control, 1995
- The PyMOL Molecular Graphics System. (2008) DeLano Scientific LLC, Palo Alto, CA, USA
- H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, *Nucleic Acids Research* 2000, 28, 235-242.
- Jones, G.; Willett, P.; Glen, R. C., Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Mol Biol* 1995, 245, (1), 43-53.